Separate Universe Approach and the Evolution of Nonlinear Superhorizon Cosmological Perturbations

G.I. Rigopoulos and E.P.S. Shellard
Department of Applied Mathematics and Theoretical Physics
Centre for Mathematical Sciences
University of Cambridge
Wilberforce Road, Cambridge, CB3 0WA, UK

In this letter we review the separate universe approach for cosmological perturbations and point out that it is essentially the lowest order approximation to a gradient expansion. Using this approach, one can study the nonlinear evolution of inhomogeneous spacetimes and find the conditions under which the long wavelength curvature perturbation can vary with time. When there is one degree of freedom or a well-defined equation of state the nonlinear long wavelength curvature perturbation remains constant. With more degrees of freedom it can vary and this variation is determined by the non-adiabatic pressure perturbation, exactly as in linear theory. We identify combinations of spatial vectors characterizing the curvature perturbation which are invariant under a change of time hypersurfaces.

I. INTRODUCTION

Until quite recently, a standard tool used in inflationary calculations was the conservation of the curvature perturbation ζ on superhorizon scales. This made inflationary predictions insensitive to what was going on between horizon crossing and horizon reentry and thus could lead to very robust predictions. However, it was realised [1] that in the presence of more degrees of freedom ζ can vary on superhorizon scales due to the presence of a non-adiabatic pressure perturbation. This fact was shown in a very simple manner in [1] using only energy momentum conservation. The rule for a varying or non-varying ζ was derived using linear perturbation theory. The extent to which the rule is true when nonlinearities are taken into account remained largely unanswered, although one would expect intuitively that it would still hold. As far as we are aware, there are three references where nonlinear conservation of a variable connected to the linear ζ has been mentioned [2, 3, 4] and, in all cases, their results refer to single field inflationary models.

Dealing with nonlinearities in general relativity is in general a very difficult issue. Yet there exists an approximation which makes the problem tractable and which is particularly relevant during inflation. The quasi-exponential expansion stretches modes to vast superhorizon scales. Hence, a long wavelength approximation for the study of inhomogeneous spacetimes smoothed on scales larger than the horizon becomes particularly relevant. The resulting picture of the inhomogeneous universe is quite simple. Each point evolves like a separate homogeneous universe with slightly different values for the Hubble rate, scalefactor, scalar field etc. One can find nonlinear variables describing the inhomogeneity which do not depend on the choice of time hypersurfaces and which are exactly gauge-invariant in the perturbation theory sense. They are combinations of spatial gradients and are immediately connected with the linear gauge invariant variables $\mathcal{R}$ and ζ (see, for example, [5] and references therein). In this letter we would like to derive such variables describing the long wavelength curvature perturbation and find their evolution equations. The latter are essentially the same as those of linear perturbation theory. In the process we also elaborate on the separate universe picture which, so far, has only been used heuristically.

II. THE LONG WAVELENGTH APPROXIMATION AND THE SEPARATE UNIVERSE PICTURE

We start by assuming that we can smooth out all relevant quantities on scales much larger than the instantaneous comoving hubble radius $\sim (aH)^{-1}$. That is, for any quantity Q we will consider the quantity

$$\bar{Q}(x) \equiv R^{-3} \int d^3x' Q(x') W(|x-x'|/R). \tag{1}$$

W is a window function rapidly decreasing for $|x - x'| > R$ and R is a smoothing scale larger than $(aH)^{-1}$. We assume that the Einstein equations act on the $\bar{Q}$'s. Of course this is not exact. Smoothing the full equations, i.e convolving the equations with the window function, is *not* obviously equivalent to the original equations acting on smoothed fields in the case of nonlinear operators. For example, if Q satisfies

$$D(Q) = 0 \tag{2}$$

then, in general,

$$\int d^3x' D(Q(x'))W(|x - x'|/R) \neq D(\bar{Q}) \tag{3}$$

if D is a nonlinear operator. Here we assume that $D(\bar{Q}) = 0$ is an adequate description of the physics at long wavelengths. Then it is a reasonable approximation to say that we can drop those terms in the equations which contain more than one spatial gradient [2]. They are expected to be small compared with time derivatives. This will be the main approximation and will lead to a picture where different points of the universe evolve independently [1].

An interesting picture emerges when one also considers the quantum fluctuations of the various physical fields during an inflationary era. In such an era the comoving horizon decreases rapidly and the various modes that are exposed as they exit the horizon can be treated as classical stochastic fields [23]. Assuming that the equations hold when acting on the $\bar{Q}$'s and writing the relevant Einstein equations as first order in time one sees that a time dependent window function results in extra terms in the equations. Indeed the time derivative of $\bar{Q}$ now has two terms,

$$\frac{\partial \bar{Q}}{\partial t} = \overline{\frac{\partial Q}{\partial t}} + \int d^3x' Q(x') \frac{\partial}{\partial t} W(|x - x'|/R) \tag{4}$$

The integration in (4) is restricted over modes close to horizon exit so $Q(x')$ can be taken to be the quantum short wavlength field which has turned classical. Hence all evolution equations get augmented by a stochastic term. This is what has been used in the past to derive an effective langevin equation for the evolution of the inflaton field. In this letter we are concerned with the evolution of the spacetime after the inhomogeneity has been set up by quntum fluctuations. We develop a stochastic framework that takes into account gravitational perturbations and is gauge invariant in the long wavelength limit in [24].

Now, consider the ADM parametrisation of the metric [6, 8]. One imagines spacetime to be foliated by spacelike hypersurfaces with a normal vector

$$n_0 = -N,\ n_i = 0,\ n^0 = N^{-1},\ n^i = N^{-1}N^i. \tag{5}$$

Here, N is called the lapse function and N^i the shift vector. The latter measures the deviation of the timelike lines which define the spatial coordinates from the integral curves of the normal to the time hypersurfaces. The metric takes the form

$$g_{00} = -N^2 + \gamma^{ij} N_i N_j\ ,\ g_{0i} = -N_i\ ,\ g_{ij} = \gamma_{ij}\ , \tag{6}$$

with the inverse

$$g^{00} = -N^{-2}\ ,\ g^{0i} = -N^{-2}N^i\ ,\ g^{ij} = \gamma^{ij} - N^{-2}N^i N^j\ , \tag{7}$$

where γ_{ij} is the metric on the spatial hypersurfaces. The four functions N and N^i parametrize the gauge freedom of general relativity and are arbitrary. The choice of gauge in the conventional language of perturbation theory is essentially the choice of N and N^i for linearized perturbations. The way the spatial hypersurfaces are embedded in the 4D geometry is parametrised by the extrinsic curvature tensor

$$K_{ij} = -\frac{1}{2N}\left(N_{i|j} + N_{j|i} + \frac{\partial}{\partial t}\gamma_{ij} \right). \tag{8}$$

Here a vertical bar denotes a covariant derivative w.r.t γ_{ij}. *From now on we will use coordinate systems with shift vector* $N^i = 0$. For matter described by the energy momentum tensor $T_{\mu\nu}$ we have the energy density

$$\mathcal{E} \equiv n^\mu n^\nu T_{\mu\nu} = N^{-2} T_{00}, \tag{9}$$

the momentum density

$$\mathcal{J}_i \equiv -n^\mu T_{\mu\, i} = -N^{-1} T_{0i}, \tag{10}$$

---

[1] Particular solutions for long wavelength equations as well as an iteration scheme based on a gradient expansion has been considered by various authors in the past, see eg. [9]. A separate universe picture has also been discussed in [10] in the context of perturbing solutions of the homogeneous equations w.r.t the constants that appear in them. Here, we are interested not in particular solutions but general properties of the nonlinear evolution of the long wavelength inhomogeneity codified by variables like (52).

and the stress tensor

$$S_{ij} \equiv T_{ij}. \tag{11}$$

A bar above a tensor will denote its traceless part

$$\bar{K}_{ij} = K_{ij} - \frac{1}{3}K\gamma_{ij}\,, \qquad K = K^i_i = \gamma^{ij}K_{ij}. \tag{12}$$

We can now write the Einstein equations for the above variables, dropping all terms which explicitely contain second order spatial derivatives and setting N^i=0. The 00 and 0i Einstein equations are

$$\bar{K}_{ij}\bar{K}^{ij} \;-\; \frac{2}{3}K^2 + \frac{16\pi}{m_{\rm pl}^2}\mathcal{E} = 0\,, \tag{13}$$

$$\bar{K}^j_{i|j} \;-\; \frac{2}{3}K_{|i} - \frac{8\pi}{m_{\rm pl}^2}\mathcal{J}_i = 0\,. \tag{14}$$

The first will turn out to be equivalent to the Friedmann equation of the FRW cosmology and the second is usually refered to as the momentum constraint. The dynamical equations are

$$\frac{dK}{dt} \;=\; N\frac{3}{2}\bar{K}_{ij}\bar{K}^{ij} + \frac{12\pi}{m_{\rm pl}^2}N\left(\mathcal{E} + \frac{1}{3}S\right)\,, \tag{15}$$

$$\frac{d\bar{K}^i_j}{dt} \;=\; N\left(K\bar{K}^i_j - \frac{8\pi}{m_{\rm pl}^2}\bar{S}^i_k\right)\,, \tag{16}$$

where $\bar{S}_{ij} = S_{ij} - \frac{1}{3}S\gamma_{ij}$ with pressure $S = S^i_i = \gamma^{ij}S_{ij}$ and $\mathcal{R}^i_j$ is the Ricci tensor of the spatial hypersurfaces. Matter will obey the continuity equations $T^{\mu\nu}{}_{;\nu} = 0$ which take the form

$$\frac{d\mathcal{E}}{dt} \;=\; NK(\mathcal{E} + \frac{1}{3}S) + N\bar{K}^{ij}\bar{S}_{ij} + N^{-1}(N^2\mathcal{J}^i)_{|i}\,, \tag{17}$$

$$\frac{d\mathcal{J}_i}{dt} \;=\; NK\mathcal{J}_i - (\mathcal{E}\delta^j_i + S^j_i)N_{|j} - NS^j_{i|j}. \tag{18}$$

In the case of matter composed of several scalar fields ϕ^A, as is relevant in scalar field driven inflation, the energy momentum tensor is

$$T_{\mu\nu} = G_{AB}\partial_\mu\phi^A\partial_\nu\phi^B - g_{\mu\nu}\left(\frac{1}{2}G_{AB}\partial^\lambda\phi^A\partial_\lambda\phi^B + V\right), \tag{19}$$

so that

$$\mathcal{E} \;\simeq\; \frac{1}{2N^2}G_{AB}\dot{\phi}^A\dot{\phi}^B + V(\phi)\,, \tag{20}$$

$$\mathcal{J}_i \;=\; -\frac{1}{N}G_{AB}\dot{\phi}^A\partial_i\phi^B\,, \tag{21}$$

$$S_{ij} \;\simeq\; \gamma_{ij}\left(\frac{1}{2N^2}G_{AB}\dot{\phi}^A\dot{\phi}^B - V\right), \tag{22}$$

where we have dropped second order spatial gradients. So in the case of a collection of scalar fields in the long wavelength approximation $(N^2\mathcal{J}_i)_{|j} \simeq 0$ and $\bar{S}_{ij} \simeq 0$, up to second order spatial gradients. Actually, for a general perfect fluid

$$T_{\mu\nu} = (\mathcal{E} + p)u_\mu u_\nu + pg_{\mu\nu}. \tag{23}$$

So, in the absence of pure vector perturbations on super-Hubble scales, $u_i = \partial_i f$ and the long wavelength approximation gives $S^i{}_j \simeq p\delta^i{}_j$, $\bar{S}^i{}_j \simeq 0$ and $\mathcal{J}_{i|j} \simeq 0$. The set of relevant evolution equations then become

$$\frac{dK}{dt} \;=\; N\frac{3}{2}\bar{K}_{ij}\bar{K}^{ij} + \frac{12\pi}{m_{\rm pl}^2}N\left(\mathcal{E} + \frac{1}{3}S\right)\,, \tag{24}$$

$$\frac{d\bar{K}^i_j}{dt} \;=\; NK\bar{K}^i_j\,, \tag{25}$$

$$\frac{d\mathcal{E}}{dt} \;=\; NK\left(\mathcal{E} + \frac{1}{3}S\right)\,, \tag{26}$$

$$\frac{d\mathcal{J}_i}{dt} \;=\; NK\mathcal{J}_i - (\mathcal{E}\delta^j_i + S^j_i)N_{|j} - NS^j_{i|j}\,, \tag{27}$$

which, along with the constraints

$$\bar{K}_{ij}\bar{K}^{ij} \;-\; \frac{2}{3}K^2 + \frac{16\pi}{m_{\rm pl}^2}\mathcal{E} = 0\,, \tag{28}$$

$$\bar{K}^j_{\;i|j} \;-\; \frac{2}{3}K_{|i} - \frac{8\pi}{m_{\rm pl}^2}\mathcal{J}_i = 0 \tag{29}$$

form the basis of this approach.

It will now be convenient to write the spatial metric as [6]

$$\gamma_{ij} = \gamma^{1/3}\tilde{\gamma}_{ij} \tag{30}$$

where γ is the determinant of the spatial metric, taking $Det(\tilde{\gamma}_{ij}) \equiv 1$. Now since $Tr(\dot{\tilde{\gamma}}_{ij}) = 0$, we have

$$\bar{K}_{ij} = -\frac{1}{2N}\gamma^{1/3}\dot{\tilde{\gamma}}_{ij} \tag{31}$$

and

$$K = -\frac{1}{2N}\frac{\dot{\gamma}}{\gamma}. \tag{32}$$

The determinant $\gamma(\mathbf{x},t)$ contains what in linear theory is usually called a scalar perturbation mode of the metric. $\tilde{\gamma}_{ij}(\mathbf{x},t)$ contains another scalar, the vector and tensor perturbations. From (32) we see that $K(\mathbf{x},t)$ can be interpreted as a locally-defined Hubble parameter since $\gamma^{1/3}(\mathbf{x},t) \equiv a(\mathbf{x},t)$ is a locally defined scalefactor. It can be explicitly shown that the system (24-29), a truncated form of Einstein's equations, forms a consistent set of equations, by which we mean that the evolution preserves the constraints. So an initial inhomogeneous configuration which respects the constraint (29), can be evolved with the evolution equations and the constraint always will be satisfied. Now observe that the only equation where spatial derivatives appear is the constraint (29) (and eqn (27) but this turns out to be irrelevant). Hence, after initial data which respect it have been specified, each point can be evolved individually. The relevant equations are essentially those of the homogeneous FRW cosmology, valid locally, modulo the $\bar{K}^i{}_j$ terms. This is exactly what has been termed in the past as the 'separate universe approach' (see, for example, [1]). It was explicitly stated in [2] in the context of general relativistic Hamilton-Jacobi theory. We see that such an approach is equivalent to the first order of a spatial gradient expansion.

Although we can solve the equations numerically retaining the $\bar{K}^i{}_j$ terms, we can readily see that they are not expected to be important dynamically [2]. Using (32) we can solve (25) to get

$$\bar{K}^i{}_j = C^i{}_j(\mathbf{x})\gamma^{-1/2}. \tag{33}$$

So, in the absence of sources, in an expanding universe (particularly for quasi-exponential expansion), the traceless part of the extrinsic curvature $\bar{K}^i{}_j$ decays rapidly. In most cases, therefore, it will be safe to ignore the $\bar{K}$ terms. Of course, in the presence of quantum fluctuations $\bar{K}$ is sourced mostly by the fluctuating γ (see (31)). Ignoring $\bar{K}$, equations (24-29) become

$$\frac{dK}{dt} \;=\; \frac{12\pi}{m_{\rm pl}^2}N\left(G_{AB}\Pi^A\Pi^B\right) \tag{34}$$

$$\frac{d}{dt}\Pi^A \;=\; NK\Pi^A - N\Gamma^A_{BC}\Pi^B\Pi^C - NG^{AB}\frac{\partial V}{\partial\phi^B} \tag{35}$$

$$K^2 \;=\; \frac{24\pi}{m_{\rm pl}^2}\left(\frac{1}{2}G_{AB}\Pi^A\Pi^B + V\right) \tag{36}$$

$$\partial_i K \;=\; \frac{12\pi}{m_{\rm pl}^2}G_{AB}\Pi^A\partial_i\phi^B, \tag{37}$$

where

$$\frac{\dot{\phi}^A}{N} = \Pi^A, \tag{38}$$

and Γ^A_{BC} is the Connection formed from G_{AB}. Eqns (34-37) are exactly the same as those of a homogeneous cosmology apart from the momentum constraint (37). Hence, the long wavelength universe looks like a collection of homogeneous

universes, each evolving according to the equations of FRW cosmology. The only information about inhomogeneity is contained in the momentum constraint (37) [2].

In single field homogeneous cosmology, it is very convenient to parametrize the Hubble rate H in terms of the value of the scalar field leading to what is called the Hamilton Jacobi formulation. A similar thing can be done here. Observe that (37) is satisfied if $K = K(\phi^A)$ and

$$\Pi^A = \frac{m_{\rm pl}^2}{12\pi} G^{AB} \frac{\partial K}{\partial \phi^B}. \tag{39}$$

Substituting this into the Friedmann equation (36) gives the Hamilton Jacobi equation for a system of scalar fields [2]

$$K^2 - \frac{m_{\rm pl}^2}{12\pi} G^{AB} \frac{\partial K}{\partial \phi^A} \frac{\partial K}{\partial \phi^B} = \frac{24\pi}{m_{\rm pl}^2} V. \tag{40}$$

So, K is a function of the scalar fields only. Note that this is consistent with equation (34) which can now be written

$$\dot{K} = \frac{\partial K}{\partial \phi^C} \dot{\phi^C}. \tag{41}$$

Hence, given appropriate boundary conditions, a solution to equation (40) determines the state of the system completely in terms of its position in the configuration space of the scalar fields. This result now holds for an inhomogeneous universe smoothed on superhorizon scales.

Solving (40) analytically is not straightforward for a general potential. Explicit solutions for the case of an exponential potential have been given in [2]. During inflation, it is usually a very good approximation to consider a slow-roll behaviour where the derivative terms can be neglected. Then (40) becomes the usual slow-roll relation

$$K^2 = \frac{24\pi}{m_{\rm pl}^2} V. \tag{42}$$

In this case the momentum constraint is equivalent to (36). Another case where an approximate solution can be found is the study of deviations of the scalar field about some stable point which are small compared to $m_{\rm pl}$ [7]. In any case, a solution to equation (40) describes the long wavelength system completely. An interesting point noted in [2], is the attractor property of solutions of (40). The latter contain a number of arbitrary parameters C^A, reflecting the freedom to choose initial conditions of the field momenta (to specify the motion one must assign not only an initial value for ϕ but also $\dot{\phi}$ at each point). Neglecting the $\bar{K}$ terms means that these parameters are spatially independent. Of course, when assigning initial conditions the momenta need not be strictly determined by the field values. So, in principle these constant parameters should be spatially dependent too. So suppose one considers a function $K(\phi^A, C)$, then

$$\begin{aligned} 2K \frac{\partial K}{\partial C} &= \frac{m_{\rm pl}^2}{6\pi} G^{AB} \frac{\partial K}{\partial \phi^A} \frac{\partial}{\partial \phi^B} \frac{\partial K}{\partial C} \\ \Rightarrow K &= \frac{m_{\rm pl}^2}{12\pi} G^{AB} \frac{\partial K}{\partial \phi^A} \frac{\partial}{\partial \phi^B} \ln \left| \frac{\partial K}{\partial C} \right| \\ &= \frac{1}{N} \frac{\partial}{\partial t} \ln \left| \frac{\partial K}{\partial C} \right| \end{aligned} \tag{43}$$

where we have used (38) and (39). From (43) and (32) we see that

$$\left| \frac{\partial K}{\partial C} \right| \propto a^{-3}. \tag{44}$$

Thus the freedom to choose the momentum independently at each point, reflected in the freedom to choose constants in the solution of (40), becomes irrelevant very quickly[2].

From now on we will also be using the more familiar notation

$$\begin{aligned} K(t, \mathbf{x}) &= -3H(t, \mathbf{x}), \\ \gamma(t, \mathbf{x}) &= a^6(t, \mathbf{x}), \end{aligned} \tag{45}$$

where a and H are the locally defined values of the scale factor and the Hubble rate.

[2] This is intimately connected to the fact that perturbation modes freeze when they exit the horizon

III. THE LONG WAVELENGTH CURVATURE PETURBATION

In linear perturbation theory, a common variable used to characterize the perturbations is $\delta\zeta$ defined as

$$\delta\zeta = \psi - H\frac{\delta\mathcal{E}}{\dot{\mathcal{E}}} = \psi - \frac{H}{\dot{H}}\left(\dot{\psi} + H\psi\right) = \psi - H\frac{\delta H}{\dot{H}} \tag{46}$$

where ψ is the scalar perturbation of the spatial part of the metric, $\psi = \delta a(t,\mathbf{x})/a(t)$. The quantity $\delta\zeta$ is gauge invariant and corresponds to the curvature perturbation in a time slicing that sets $\delta\mathcal{E} = 0$. In the case of single field inflation, where there is one degree of freedom, on long wavelengths, $\delta\dot{\zeta} = 0$. When there are many scalar fields present [1]

$$\delta\dot{\zeta} = -\frac{H}{(\mathcal{E}+p)}\delta p_{nad}, \tag{47}$$

where δp_{nad} is the non-adiabatic pressure perturbation. In general, when there is a well defined equation of state $p = p(\mathcal{E})$, δp_{nad}=0. We will now derive a nonlinear generalisation of $\delta\zeta$ and eqn (47) for the nonlinear long wavelength case.

Consider a generic time slicing and some initial time t_I. We can integrate the local metric determinant up to a surface of constant energy density using (32)

$$\ln a(t_I,\mathbf{x}) = \ln a(t,\mathbf{x}) - \int_{t_I}^{T} N(t',\mathbf{x})H(t',\mathbf{x})dt', \tag{48}$$

where $T\left(\mathcal{E}(\mathbf{x})\right)$ is the value of t at the point $\mathbf{x}$ for $\mathcal{E}(\mathbf{x}) = const$. Taking the spatial derivative and using (32) once more we see that

$$\begin{aligned} \partial_i\left[\ln a(t_I,\mathbf{x})\right] &= \left(\partial_i - \frac{\partial T}{\partial x^i}\partial_t\right)\ln a(t,\mathbf{x}) - \int_{t_I}^{T}\partial_i\left[N(t',\mathbf{x})H(t',\mathbf{x})\right]dt' \\ &= \left(\partial_i - \frac{\partial T}{\partial x^i}\partial_t\right)\ln\gamma(t,\mathbf{x}) - \int_{t_I}^{T}\partial_i\partial_t\left[\ln a(t',\mathbf{x})\right]\,dt' \end{aligned} \tag{49}$$

which can be written

$$\partial_i[\ln a(T,\mathbf{x})] = \partial_i[\ln a(t,\mathbf{x})] - \frac{\partial T}{\partial x^i}N(t,\mathbf{x})H(t,\mathbf{x}), \tag{50}$$

or

$$X_i^{(\mathcal{E})} = X_i - \frac{NH}{d\mathcal{E}/dT}\partial_i\mathcal{E}, \tag{51}$$

where we have defined $X_i \equiv \partial_i \ln a$ and $T = T(\mathcal{E})$. Note that the gradient of the l.h.s of (50) is evaluated along the hypersurface $\mathcal{E} = const$. We see that with any choice of N, the particular combination of gradients appearing on the r.h.s of (51) always equals the gradient of $\ln a$ on a surface where $\mathcal{E} = const$. So we are led to examine the variable

$$\zeta_i = X_i - \frac{NH}{\dot{\mathcal{E}}}\partial_i\mathcal{E}, \tag{52}$$

as a coordinate independent measure of the nonlinear curvature perturbation. Let us first see how this variable evolves with time. Taking the time derivative of (52) and using (26) we arrive at

$$\dot{\zeta}_i = \frac{1}{3}\frac{1}{\left(\mathcal{E}+\frac{1}{3}S\right)^2}\left(\dot{\mathcal{E}}\partial_i S - \dot{S}\partial_i\mathcal{E}\right), \tag{53}$$

or, by using (26) once more

$$\dot{\zeta}_i = -\frac{1}{3}\frac{NH}{\left(\mathcal{E}+\frac{1}{3}S\right)}\left(\partial_i S - \frac{\dot{S}}{\dot{\mathcal{E}}}\partial_i\mathcal{E}\right). \tag{54}$$

This should be compared with (47). It is easy to see that if there is a well defined equation of state $S = S(\mathcal{E})$, then the r.h.s of (53) is zero. Therefore, the long wavelength curvature perturbation, described by the vector ζ_i, can evolve on superhorizon scales if and only if there exists a non-adiabatic pressure perturbation. This was proved in [1] for the case of linear perturbation theory by making use only of energy conservation. The same applied here. Hence equation (28) shows that this result extends to the nonlinear case if we use the vector ζ_i to describe the perturbation.

The connection with the usual linearized gauge invariant variable ζ is obvious from (52). Writing $X = \ln(\gamma)$, $\mathcal{E} = \mathcal{E}(t) + \delta\mathcal{E}$, $\gamma = \gamma(t) + \delta\gamma$ and dropping terms quadratic in the perturbations, we see that

$$\begin{aligned} \zeta_i &\simeq \partial_i \delta\zeta\,, \\ \delta\zeta &\simeq \frac{\delta a}{a} - \left(\frac{NH}{\dot{\mathcal{E}}}\right)(t, \mathbf{x}_0)\, \delta\mathcal{E} + C(t), \end{aligned} \tag{55}$$

with C(t) the background number of e-folds.

The quantity ζ_i was defined with respect to the energy density. Also the way we arrived at it was at best heuristic. We will now show that for any inhomogeneous spacetime scalar $\phi(t, \mathbf{x})$, the variable

$$\mathcal{R}_i = X_i - \frac{NH}{\dot{\phi}} \partial_i \phi \tag{56}$$

is gauge invariant in the long wavelength approximation under mild conditions regarding the allowed coordinate transformations. Consider the transformation

$$(t, \mathbf{x}) \to (T(t, \mathbf{x}), \mathbf{X}(t, \mathbf{x}))\,. \tag{57}$$

Then, it can be shown that [2]

$$X^j = f^j(x^i) + \int \frac{T^{,j}}{T^{,0} T_{,0}} dT, \tag{58}$$

where f^j is an arbitrary function independent of time. One can choose it for convenience to be $\delta^i_j x^j$ [2] but one can also attach a physical meaning to such a choice as can be seen in figure 1. Related considerations in matching perturbed and homogeneous spacetimes are discussed in [11]. We note that the transformation (58) holds between gauges with the shift $N_i = 0$. Assuming that the new time coordinate T is non-singular, that is, that the new time hypersurfaces do not get too 'wrinkly', we can discard the gradient of the integral as second order in spatial gradients. We then have

$$\frac{\partial X^j}{\partial x^i} \simeq \delta^j{}_i. \tag{59}$$

Consider now a gauge with coordinates x^μ where $\partial_i \phi = 0$ and a transformation to a different gauge with coordinates $X^{\tilde{\mu}}$. Then

$$\begin{aligned} \gamma_{kl} &= \frac{\partial X^{\tilde{\mu}}}{\partial x^k} \frac{\partial X^{\tilde{\nu}}}{\partial x^l} g_{\tilde{\mu}\tilde{\nu}} \\ &\simeq -T_{,k} T_{,l} N_T^2 + \delta^{\tilde{k}}{}_k \delta^{\tilde{l}}{}_l \gamma_{\tilde{k}\tilde{l}} \\ &\simeq \delta^{\tilde{k}}{}_k \delta^{\tilde{l}}{}_l \gamma_{\tilde{k}\tilde{l}}\,, \end{aligned} \tag{60}$$

and so

$$\gamma(t, \mathbf{x}) \simeq \tilde{\gamma}(T, \mathbf{X}). \tag{61}$$

Taking the spatial derivative w.r.t x^i we get

$$\begin{aligned} \frac{\partial}{\partial x^i} \ln \gamma &= \frac{\partial T}{\partial x^i} \partial_T \ln \tilde{\gamma} + \frac{\partial X^j}{\partial x^i} \frac{\partial}{\partial X^j} \ln \tilde{\gamma} \\ &\simeq \partial_i T \partial_T \ln \tilde{\gamma} + \frac{\partial}{\partial X^i} \ln \tilde{\gamma}. \end{aligned} \tag{62}$$

For the scalar $\phi(t, \mathbf{x})$ we have $\partial_i \phi = 0$, so in the new coordinates

$$\partial_i T \partial_T \tilde{\phi}(T, \mathbf{X}) + \frac{\partial}{\partial X^i} \tilde{\phi}(T, \mathbf{X}) = 0. \tag{63}$$

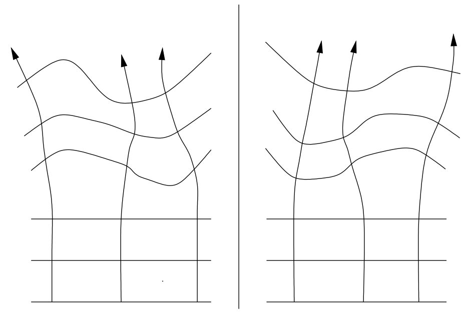

FIG. 1: Two different slicings of an inhomogeneous spacetime based on two different time variables, T and $\tilde{T}$, have different normal curves defining the spatial coordinates. A particular patch of spacetime during inflation starts its life inside the horizon where it can be considered as homogeneous. Classical perturbations are generated when modes cross the horizon and freeze in. While subhorizon and homogeneous, there is a preferred time slicing in the patch to which all slicings and spatial coordinate choices should match. This fixes $f^j(x^i) = \delta^j_i x^i$. This choice separates changes in the time slicing – which we want to consider as a gauge choice – from possible coordinate transformations of the homogeneous spacetime.

Therefore

$$\partial_i \ln\gamma = \frac{\partial}{\partial X^i}\ln\tilde{\gamma} - \frac{\partial_T \ln\tilde{\gamma}}{\partial_T\phi}\frac{\partial}{\partial X^i}\phi\,. \tag{64}$$

Now, since the transformation we used was arbitrary (up to considerations for the smooth behaviour of the new time T), and noting that

$$H = -\frac{1}{6N_T}\partial_T \ln\tilde{\gamma} \tag{65}$$

and $X_i = \partial_i \ln a$, we conclude that the variable

$$\mathcal{R}_i = X_i - \frac{NH}{\dot{\phi}}\partial_i\phi, \tag{66}$$

equals the gradient of the spatial metric determinant in a gauge where the scalar ϕ is homogeneous; it is, therefore, a nonlinear gauge invariant measure of the perturbations. When the scalar is taken to be the energy density $\mathcal{E}$ we recover ζ_i of eqn (52).

We would like to remark here that the use of spatial gradients to describe perturbations in a gauge invariant manner was first advocated in [12]. There, the use of the word 'gauge' refers to a choice of a correspondence between a fictitious background homogeneous spacetime and the real perturbed universe. So a gauge invariant quantity is one that does not change when this correspondence is altered, which is achieved when the quantity either vanishes or is a trivial tensor in the background [13, 14]. In this sense, any gradient is gauge invariant since it must vanish on the homogeneous background. Here, however, by a choice of gauge we mean a choice of the lapse function N and the shift vector N_i [5], so we look for quantities such as (66) that are invariant under a change of time slicing as well.

Examples

When the scalar ϕ in (56) is taken to be the inflaton it can be easily verifield that $\mathcal{R}_i$ is also conserved. Using eqns (34–37) one can show that

$$\dot{\mathcal{R}}_i = 0\,. \tag{67}$$

When there is more than one scalar field present, however, the situation is different. It is well known in linear perturbation theory [1, 15, 16] that perturbations in multi-scalar field models can source the curvature perturbation $\delta\zeta$ on long wavelengths, making the time evolution of the latter possible. Let us see how this effect is very easily seen in the formalism developed in the previous section without the need to resort to linear perturbation theory. Consider hypersurfaces with $\partial_i\mathcal{E} = 0$. From (28) with the $\bar{K}$ terms omitted we see that

$$\mathcal{E}_{,i} = 0 \Rightarrow H_{,i} = 0 \Rightarrow \mathcal{J}_i = 0 \Rightarrow G_{AB}\Pi^A\partial_i\phi^B = 0. \tag{68}$$

The fact that $\mathcal{E}_{,i} = 0$ also means that

$$S_{,i} = 2G_{AB}\Pi^A\partial_i\Pi^B. \tag{69}$$

So, if there is a single field,

$$\mathcal{E}_{,i} = 0 \;\Rightarrow\; H_{,i} = 0 \;\Rightarrow\; \partial_i \phi = 0 \;\Rightarrow\; S_{,i} = 0\,, \tag{70}$$

since $\partial_i \Pi = 0$. Hence ζ_i remains constant. On the other hand, for many fields, we have

$$\mathcal{E}_{,i} = 0 \;\Rightarrow\; H_{,i} = 0 \;\Rightarrow\; G_{AB}\Pi^A \partial_i \phi^B = 0. \tag{71}$$

So, generally in this case, we may have $\partial_i \phi^A \neq 0$. Therefore, from (69) $\partial_i S \neq 0$ in general and ζ_i can evolve on superhorizon scales, $\dot{\zeta}_i \neq 0$.

In general, for a system of scalar fields governed by the Lagrangian

$$\mathcal{L} = \frac{1}{2} G_{AB} \partial^\mu \phi^A \partial_\mu \phi^B - V, \tag{72}$$

the variation of ζ_i from (54) is given by

$$\dot{\zeta}_i = V_A \left[\frac{d}{dt} \left(\frac{1}{G_{BC}\Pi^B \Pi^C} \right) \partial_i \phi^A - \dot{\phi}^A \partial_i \left(\frac{1}{G_{BC}\Pi^B \Pi^C} \right) \right]. \tag{73}$$

In the multi field case it is convenient to consider the variables

$$\mathcal{Q}_i^A = \Pi^A X_i - H \partial_i \phi^A, \tag{74}$$

which are also invariant under changes of time hypersurfaces in the long wavelength limit. They are not conserved even in the case of a single field. In particular

$$\dot{\mathcal{Q}}_i^A = \dot{\Pi}^A \mathcal{R}_i^A + \Pi^A \dot{\mathcal{R}}_i^A, \tag{75}$$

where no summation is implied. It should be noted here that it is possible to define conserved variables for each field if and only if the potential is of the form $\sum V\left(\phi^A\right)$ or if one field dominates completely over the other. In terms of the $\mathcal{Q}_i^A$'s the curvature perturbation is given by

$$\zeta_i = \frac{G_{AB}\Pi^A \mathcal{Q}_i^B}{G_{CD}\Pi^C \Pi^D}. \tag{76}$$

Consider now the following simple two-field example which can be considered a toy model for hybrid inflation. Initially the first field σ is stuck slightly displaced from the bottom of a valley while the other field ϕ evolves and drives inflation. In this case, we have essentially a single field model and ζ_i is conserved. With $\Pi^\sigma \simeq 0$, we have

$$\zeta_i^{(I)} = X_i - \frac{H}{\Pi^\phi} \left(\partial_i \phi + \frac{G_{\phi\sigma}}{G_{\phi\phi}} \partial_i \sigma \right) . \tag{77}$$

When the ϕ field reaches a critical value, the trajectory turns abruptly in the direction of variations in σ which can now dominate until inflation ends. For the period of rolling in the σ-direction we have an extra contribution to ζ_i

$$\zeta_i^{(II)} = X_i - \frac{H}{\Pi^\sigma} \left(\partial_i \sigma + \frac{G_{\phi\sigma}}{G_{\sigma\sigma}} \partial_i \phi \right) . \tag{78}$$

Note that for a non-diagonal metric G_{AB} the isocurvature σ-field can contribute to the curvature perturbation even for a straight trajectory [16, 17], in contrast to what happens when $G_{AB} = \delta_{AB}$ [15]. Also, if during the rolling along the σ direction we have a second period of inflation, the σ field can contribute to the curvature perturbation since

$$\zeta_i = \zeta_i^{(I)} + \zeta_i^{(II)}. \tag{79}$$

In the case with $G_{AB} = \delta_{AB}$, we can easily find the following evolution equations for the $\mathcal{R}_i$'s in the case of two fields

$$\dot{\mathcal{R}}_i^\phi = \frac{4\pi}{m_{\rm pl}^2} N \Pi_\sigma \left(\partial_i \sigma - \frac{\Pi_\sigma}{\Pi_\phi} \partial_i \phi \right) \tag{80}$$

and

$$\dot{\mathcal{R}}_i^\sigma = \frac{4\pi}{m_{\rm pl}^2} N \Pi_\phi \left(\partial_i \phi - \frac{\Pi_\phi}{\Pi_\sigma} \partial_i \sigma \right) . \tag{81}$$

For many general fluids one can write the curvature perturbation as a weighted sum over components [25]

$$\zeta_i = \frac{\sum \dot{\mathcal{E}}_A \zeta_i^A}{\sum \dot{\mathcal{E}}_A}. \tag{82}$$

If the components are non interacting, each will obey an evolution equation like (26), so each ζ_i^A is separately conserved. Hence one can see that the time evolution of the total curvature perturbation is determined by the ratio of the energy densities of the various components. For example, starting with $\zeta_i^A = 0$ we get

$$\zeta_i = \frac{\dot{\mathcal{E}}_B}{\dot{\mathcal{E}}_A + \dot{\mathcal{E}}_B} \zeta_i^B \tag{83}$$

This evolution of the long wavelength curvature perturbation in multifield models is one of the possible mechanisms that could potentially produce significant non gaussianity. Perturbations from the σ direction can be significantly non-gaussian and now there is the possibility that these fluctuations can contribute to the curvature perturbation at later times [18, 19, 20]. This can be seen for example from equations (80) and (82). Other mechanisms relating non-gaussianity with a second scalar field have been proposed [21]. We will return on the issue with more details in [24].

One can also easily see that there is a particular form of a multi-scalar field potential which will conserve ζ_i. With $H_{,i} = 0$ we have $G_{AB}\Pi^A \partial_i \phi^B = 0$, so to get $S_{,i} = 0$ we need (see (69) and (39))

$$G_{AB}\Pi^A \partial_i \Pi^B = \frac{m_{\rm pl}^2}{12\pi} G_{AB}\Pi^A \partial_i \left(G^{BC} \frac{\partial H}{\partial \phi^C} \right) = 0\,. \tag{84}$$

Thus, if

$$G^{AC} \frac{\partial H}{\partial \phi^C} = \frac{\lambda}{m_{\rm pl}} \phi^A \;\Rightarrow\; \frac{\partial H}{\partial \phi^A} = \frac{\lambda}{m_{\rm pl}} G_{AB} \phi^B, \tag{85}$$

with λ some dimensionless number, then $S_{,i} = 0$. From (40)

$$H^2 = \frac{m_{\rm pl}^2}{12\pi} G^{AB} \frac{\partial H}{\partial \phi^A} \frac{\partial H}{\partial \phi^B} + \frac{8\pi}{3m_{\rm pl}^2} V(\phi). \tag{86}$$

So, with

$$H = \frac{\lambda}{m_{\rm pl}} \frac{1}{2} G_{AB} \phi^A \phi^B + \mu m_{\rm pl}, \tag{87}$$

we get

$$V(\phi) = -\frac{1}{2} M^2 \phi^A \phi_A + M^2 \frac{\pi}{2m_{\rm pl}^2} \left(\frac{1}{2} \phi^A \phi_A + q m_{\rm pl}^2 \right)^2, \tag{88}$$

with q another dimensionless parameter and M^2 can have either sign. So, for a general potential containing both quartic and quadratic terms with the various constants related as in (76), the long wavelength curvature perturbation does not evolve in time. One should keep in mind that the above results only apply when anisotropic stresses are negligible on superhorizon scales and therefore one can safely neglect the $\bar{K}^i_j$ terms.

We emphasise again that for the derivation of all of the above no linear approximation has been made. The equations are exact and hold at every point of the inhomogeneous spacetime in the long wavelength limit. For more uses of this formalism we refer the reader to [24].

IV. SUMMARY

We have shown that particular combinations of spatial gradients are good variables to describe the curvature perturbation in a long wavelength limit. We also elaborated on the intuitive picture of the 'separate universe approach' where, given initial data which respect the energy and momentum constraints, one can view the long wavelength universe as a collection of universes each evolving independently. Such a picture is a simple way to understand

nonlinear evolution on large scales and, in particular, to determine the conditions under which the long wavelength curvature perturbation can vary. We believe this is a remarkably straightforward and physically intuitive approach which should find much wider application.

Note added: While this paper was undergoing final revision, independent work by Lyth and Wands appeared on the archive [25]. This addresses some related issues on ζ-conservation using linear perturbation methods.

[1] D.Wands, K.A. Malik, D.H. Lyth, A.R. Liddle, Phys. Rev. D 62, (2000) 043527.
[2] D.S. Salopek, J.R. Bond, Phys. Rev. D43:1005-1031,1991,ibid. Phys.Rev.D 42 (1990) 3936.
[3] N. Afshordi, R. Brandenberger, Rhys.Rev.D 63, 123505.
[4] J. Maldacena, astro-ph/0210603.
[5] A. Liddle, D.H. Lyth, "Cosmological Inflation and Large Scale Structure", CUP 2000.
[6] J. M. Bardeen, Phys. Rev. D 22, (1980) 1882.
[7] W.H. Kinney, Phys. Rev. D 56, (1997) 2002.
[8] Misner, Thorne and Wheeler, 'Gravitation'.
[9] G.L. Comer et. al, Phys.Rev.D 49, 2759(1994),
N. Deruelle & D. Langlois, Phys.Rev.D 52, 2007(1995) and references therein,
J. Parry et. al. Phys.Rev.D 49, 2872(1994)
I.M. Khalatnikov et. al, Class.Quant.Grav. 19(2002) 3845
[10] M. Sasaki & T. Tanaka, Prog.Theor.Phys.99:763-782,1998,
Y. Nambu & A. Taruya, Class.Quant.Grav. 15(1998) 2761,
H. Kodama & T. Hamazaki, Phys.Rev.D 57 7177 (1998)
[11] G. Amery & E.P.S. Shellard, Phys. Rev. D67 (2003), 083502,
J. Uzan, N. Deruelle & N. Turok, Phys.Rev. D57 (1998) 7192-7199.
[12] G.F.R. Ellis, M. Bruni, Phys. Rev. D, 40 (1989), 1804.
[13] J.M. Stewart, Class. Quant. Grav. 7 (1990) 1169-1180.
[14] M. Bruni, S. Sonego, Class. Quant. Grav. 16(1999) L29-L36.
[15] C. Gordon, D. Wands, B.A. Bassett, R. Maartens, Phys. Rev. D 63 (2001) 023506 .
[16] G. Rigopoulos, astro-ph/0212141.
[17] F. DiMarco, F. Finelli, R. Brandenberger, astro-ph/0211276.
[18] A. Linde & V. Mukhanov, Phys.Rev.D56:535-539,1997
[19] F. Bernardeau, J. Uzan, Phys. Rev. D66 (2002) 103506.
[20] D.H. Lyth, C. Ungarelli, D. Wands, Phys. Rev. D67 (2003) 023503.
[21] M. Zaldarriaga, astro-ph/0306006.
[22] V. Acquaviva, N Bartolo, S. Matarrese, A. Riotto, astro-ph/020915.
[23] A. Starobinsky, in *De Vega, H.j. (Ed.), Sanchez, N. (Ed.): Field Theory, Quantum Gravity and Strings*, 107-126, 1986.
[24] G.I. Rigopoulos and E.P.S. Shellard, in preparation.
[25] D. Lyth & D. Wands, astro-ph/0306498.